Toni Scheiner

RAW SEQUENCE LISTING
PATENT APPLICATION *US/08/846,017*

DATE: 03/04/98
TIME: 18:53:44

INPUT SET: S24002.raw

This Raw Listing contains the General Information Section and up to the first 5 pages.

#10 Amh 6-4-98

ENTERED

SEQUENCE LISTING

(1) General Information:

(i) APPLICANT: Cech, Thomas R.
Lingner, Joachim
Nakamura, Toru
Chapman, Karen B.
Morin, Gregg B.
Harley, Calvin
Andrews, William H.

(ii) TITLE OF INVENTION: Novel Telomerase

(iii) NUMBER OF SEQUENCES: 171

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESS: Townsend and Townsend and Crew LLP
(B) STREET: Two Embarcadero Center, 8th Floor
(C) CITY: San Francisco
(D) STATE: California
(E) COUNTRY: United States of America
(F) ZIP: 94111

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: US 08/846,017
(B) FILING DATE: 25-APR-1997
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 08/844,419
(B) FILING DATE: 18-APR-1997
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 08/724,643
(B) FILING DATE: 01-OCT-1996
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Apple, Randolph T.
(B) REGISTRATION NUMBER: 36,429
(C) REFERENCE/DOCKET NUMBER: 015389-002920US

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (415) 576-0200
(B) TELEFAX: (415) 576-0300

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3279 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

AAAACCCCAA AACCCCAAAA CCCCTTTTAG AGCCCTGCAG TTGGAAATAT AACCTCAGTA 60

TTAATAAGCT CAGATTTTAA ATATTAATTA CAAAACCTAA ATGGAGGTTG ATGTTGATAA 120

TCAAGCTGAT AATCATGGCA TTCACTCAGC TCTTAAGACT TGTGAAGAAA TTAAAGAAGC 180

TAAAACGTTG TACTCTTGGA TCCAGAAAGT TATTAGATGA AGAAATCAAT CTCAAAGTCA 240

TTATAAAGAT TTAGAAGATA TTAAAATATT TGCGCAGACA AATATTGTTG CTACTCCACG 300

AGACTATAAT GAAGAAGATT TTAAAGTTAT TGCAAGAAAA GAAGTATTTT CAACTGGACT 360

AATGATCGAA CTTATTGACA AATGCTTAGT TGAACTTCTT TCATCAAGCG ATGTTTCAGA 420

TAGACAAAAA CTTCAATGAT TTGGATTTCA ACTTAAGGGA AATCAATTAG CAAAGACCCA 480

TTTATTAACA GCTCTTTCAA CTCAAAAGCA GTATTTCTTT CAAGACGAAT GGAACCAAGT 540

TAGAGCAATG ATTGGAAATG AGCTCTTCCG ACATCTCTAC ACTAAATATT TAATATTCCA 600

GCGAACTTCT GAAGGAACTC TTGTTCAATT TTGCGGGAAT AACGTTTTTG ATCATTTGAA 660

AGTCAACGAT AAGTTTGACA AAAAGCAAAA AGGTGGAGCA GCAGACATGA ATGAACCTCG 720

ATGTTGATCA ACCTGCAAAT ACAATGTCAA GAATGAGAAA GATCACTTTC TCAACAACAT 780

CAACGTGCCG AATTGGAATA ATATGAAATC AAGAACCAGA ATATTTTATT GCACTCATTT 840

TAATAGAAAT AACCAATTCT TCAAAAAGCA TGAGTTTGTG AGTAACAAAA ACAATATTTC 900

AGCGATGGAC AGAGCTCAGA CGATATTCAC GAATATATTC AGATTTAATA GAATTAGAAA 960

100 GAAGCTAAAA GATAAGGTTA TCGAAAAAAT TGCCTACATG CTTGAGAAAG TCAAAGATTT 1020
101
102 TAACTTCAAC TACTATTTAA CAAAATCTTG TCCTCTTCCA GAAAATTGGC GGGAACGGAA 1080
103
104 ACAAAAAATC GAAAACTTGA TAAATAAAAC TAGAGAAGAA AAGTCGAAGT ACTATGAAGA 1140
105
106 GCTGTTTAGC TACACAACTG ATAATAAATG CGTCACACAA TTTATTAATG AATTTTTCTA 1200
107
108 CAATATACTC CCCAAAGACT TTTTGACTGG AAGAAACCGT AAGAATTTTC AAAAGAAAGT 1260
109
110 TAAGAAATAT GTGGAACTAA ACAAGCATGA ACTCATTCAC AAAAACTTAT TGCTTGAGAA 1320
111
112 GATCAATACA AGAGAAATAT CATGGATGCA GGTTGAGACC TCTGCAAAGC ATTTTTATTA 1380
113
114 TTTTGATCAC GAAAACATCT ACGTCTTATG GAAATTGCTC CGATGGATAT TCGAGGATCT 1440
115
116 CGTCGTCTCG CTGATTAGAT GATTTTTCTA TGTCACCGAG CAACAGAAAA GTTACTCCAA 1500
117
118 AACCTATTAC TACAGAAAGA ATATTTGGGA CGTCATTATG AAAATGTCAA TCGCAGACTT 1560
119
120 AAAGAAGGAA ACGCTTGCTG AGGTCCAAGA AAAAGAGGTT GAAGAATGGA AAAAGTCGCT 1620
121
122 TGGATTTGCA CCTGGAAAAC TCAGACTAAT ACCGAAGAAA ACTACTTTCC GTCCAATTAT 1680
123
124 GACTTTCAAT AAGAAGATTG TAAATTCAGA CCGGAAGACT ACAAAATTAA CTACAAATAC 1740
125
126 GAAGTTATTG AACTCTCACT TAATGCTTAA GACATTGAAG AATAGAATGT TTAAAGATCC 1800
127
128 TTTTGGATTC GCTGTTTTTA ACTATGATGA TGTAATGAAA AAGTATGAGG AGTTTGTTTG 1860
129
130 CAAATGGAAG CAAGTTGGAC AACCAAAACT CTTCTTTGCA ACTATGGATA TCGAAAAGTG 1920
131
132 ATATGATAGT GTAAACAGAG AAAAACTATC AACATTCCTA AAAACTACTA AATTACTTTC 1980
133
134 TTCAGATTTC TGGATTATGA CTGCACAAAT TCTAAAGAGA AAGAATAACA TAGTTATCGA 2040
135
136 TTCGAAAAAC TTTAGAAAGA AAGAAATGAA AGATTATTTT AGACAGAAAT TCCAGAAGAT 2100
137
138 TGCACTTGAA GGAGGACAAT ATCCAACCTT ATTCAGTGTT CTTGAAAATG AACAAAATGA 2160
139
140 CTTAAATGCA AAGAAAACAT TAATTGTTGA AGCAAAGCAA AGAAATTATT TTAAGAAAGA 2220
141
142 TAACTTACTT CAACCAGTCA TTAATATTTG CCAATATAAT TACATTAACT TTAATGGGAA 2280
143
144 GTTTTATAAA CAAACAAAAG GAATTCCTCA AGGTCTTTGA GTTTCATCAA TTTTGTCATC 2340
145
146 ATTTTATTAT GCAACATTAG AGGAAAGCTC CTTAGGATTC CTTAGAGATG AATCAATGAA 2400
147
148 CCCTGAAAAT CCAAATGTTA ATCTTCTAAT GAGACTTACA GATGACTATC TTTTGATTAC 2460
149
150 AACTCAAGAG AATAATGCAG TATTGTTTAT TGAGAAACTT ATAAACGTAA GTCGTGAAAA 2520
151
152 TGGATTTAAA TTCAATATGA AGAAACTACA GACTAGTTTT CCATTAAGTC CAAGCAAATT 2580

153
154 TGCAAAATAC GGAATGGATA GTGTTGAGGA GCAAAATATT GTTCAAGATT ACTGCGATTG 2640
155
156 GATTGGCATC TCAATTGATA TGAAAACTCT TGCTTTAATG CCAAATATTA ACTTGAGAAT 2700
157
158 AGAAGGAATT CTGTGTACAC TCAATCTAAA CATGCAAACA AAGAAAGCAT CAATGTGGCT 2760
159
160 CAAGAAGAAA CTAAAGTCGT TTTTAATGAA TAACATTACC CATTATTTTA GAAAGACGAT 2820
161
162 TACAACCGAA GACTTTGCGA ATAAAACTCT CAACAAGTTA TTTATATCAG GCGGTTACAA 2880
163
164 ATACATGCAA TGAGCCAAAG AATACAAGGA CCACTTTAAG AAGAACTTAG CTATGAGCAG 2940
165
166 TATGATCGAC TTAGAGGTAT CTAAAATTAT ATACTCTGTA ACCAGAGCAT TCTTTAAATA 3000
167
168 CCTTGTGTGC AATATTAAGG ATACAATTTT TGGAGAGGAG CATTATCCAG ACTTTTTCCT 3060
169
170 TAGCACACTG AAGCACTTTA TTGAAATATT CAGCACAAAA AAGTACATTT TCAACAGAGT 3120
171
172 TTGCATGATC CTCAAGGCAA AAGAAGCAAA GCTAAAAAGT GACCAATGTC AATCTCTAAT 3180
173
174 TCAATATGAT GCATAGTCGA CTATTCTAAC TTATTTTGGA AAGTTAATTT TCAATTTTTG 3240
175
176 TCTTATATAC TGGGGTTTTG GGGTTTTGGG GTTTTGGGG 3279
177
178 (2) INFORMATION FOR SEQ ID NO:2:
179
180 (i) SEQUENCE CHARACTERISTICS:
181 (A) LENGTH: 1031 amino acids
182 (B) TYPE: amino acid
183 (C) STRANDEDNESS: Not Relevant
184 (D) TOPOLOGY: Not Relevant
185
186 (ii) MOLECULE TYPE: protein
187
188 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:
189
190 Met Glu Val Asp Val Asp Asn Gln Ala Asp Asn His Gly Ile His Ser
191 1 5 10 15
192
193 Ala Leu Lys Thr Cys Glu Glu Ile Lys Glu Ala Lys Thr Leu Tyr Ser
194 20 25 30
195
196 Trp Ile Gln Lys Val Ile Arg Cys Arg Asn Gln Ser Gln Ser His Tyr
197 35 40 45
198
199 Lys Asp Leu Glu Asp Ile Lys Ile Phe Ala Gln Thr Asn Ile Val Ala
200 50 55 60
201
202 Thr Pro Arg Asp Tyr Asn Glu Glu Asp Phe Lys Val Ile Ala Arg Lys
203 65 70 75 80
204
205 Glu Val Phe Ser Thr Gly Leu Met Ile Glu Leu Ile Asp Lys Cys Leu

```
                    85                  90                  95

Val Glu Leu Leu Ser Ser Ser Asp Val Ser Asp Arg Gln Lys Leu Gln
            100                 105                 110

Cys Phe Gly Phe Gln Leu Lys Gly Asn Gln Leu Ala Lys Thr His Leu
        115                 120                 125

Leu Thr Ala Leu Ser Thr Gln Lys Gln Tyr Phe Phe Gln Asp Glu Trp
    130                 135                 140

Asn Gln Val Arg Ala Met Ile Gly Asn Glu Leu Phe Arg His Leu Tyr
145                 150                 155                 160

Thr Lys Tyr Leu Ile Phe Gln Arg Thr Ser Glu Gly Thr Leu Val Gln
                165                 170                 175

Phe Cys Gly Asn Asn Val Phe Asp His Leu Lys Val Asn Asp Lys Phe
            180                 185                 190

Asp Lys Lys Gln Lys Gly Gly Ala Ala Asp Met Asn Glu Pro Arg Cys
        195                 200                 205

Cys Ser Thr Cys Lys Tyr Asn Val Lys Asn Glu Lys Asp His Phe Leu
    210                 215                 220

Asn Asn Ile Asn Val Pro Asn Trp Asn Asn Met Lys Ser Arg Thr Arg
225                 230                 235                 240

Ile Phe Tyr Cys Thr His Phe Asn Arg Asn Asn Gln Phe Phe Lys Lys
                245                 250                 255

His Glu Phe Val Ser Asn Lys Asn Asn Ile Ser Ala Met Asp Arg Ala
            260                 265                 270

Gln Thr Ile Phe Thr Asn Ile Phe Arg Phe Asn Arg Ile Arg Lys Lys
        275                 280                 285

Leu Lys Asp Lys Val Ile Glu Lys Ile Ala Tyr Met Leu Glu Lys Val
    290                 295                 300

Lys Asp Phe Asn Phe Asn Tyr Tyr Leu Thr Lys Ser Cys Pro Leu Pro
305                 310                 315                 320

Glu Asn Trp Arg Glu Arg Lys Gln Lys Ile Glu Asn Leu Ile Asn Lys
                325                 330                 335

Thr Arg Glu Glu Lys Ser Lys Tyr Tyr Glu Glu Leu Phe Ser Tyr Thr
            340                 345                 350

Thr Asp Asn Lys Cys Val Thr Gln Phe Ile Asn Glu Phe Phe Tyr Asn
        355                 360                 365
```

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION *US/08/846,017*

DATE: 03/04/98
TIME: 18:54:03

INPUT SET: S24002.raw

Line	Error	Original Text

Application No.: 08/846017

NOTICE TO COMPLY WITH REQUIREMENTS FOR PATENT APPLICATIONS CONTAINING NUCLEOTIDE SEQUENCE AND/OR AMINO ACID SEQUENCE DISCLOSURES

The nucleotide and/or amino acid sequence disclosure contained in this application does not comply with the requirements for such a disclosure as set forth in 37 C.F.R. 1.821 - 1.825 for the following reason(s):

☒ 1. This application clearly fails to comply with the requirements of 37 C.F.R. 1.821-1.825. Applicant's attention is directed to these regulations, published at 1114 OG 29, May 15, 1990 and at 55 FR 18230, May 1, 1990.

☐ 2. This application does not contain, as a separate part of the disclosure on paper copy, a "Sequence Listing" as required by 37 C.F.R. 1.821(c).

☐ 3. A copy of the "Sequence Listing" in computer readable form has not been submitted as required by 37 C.F.R. 1.821(e).

☐ 4. A copy of the "Sequence Listing" in computer readable form has been submitted. However, the content of the computer readable form does not comply with the requirements of 37 C.F.R. 1.822 and/or 1.823, as indicated on the attached copy of the marked -up "Raw Sequence Listing."

☐ 5. The computer readable form that has been filed with this application has been found to be damaged and/or unreadable as indicated on the attached CRF Diskette Problem Report. A Substitute computer readable form must be submitted as required by 37 C.F.R. 1.825(d).

☐ 6. The paper copy of the "Sequence Listing" is not the same as the computer readable from of the "Sequence Listing" as required by 37 C.F.R. 1.821(e).

☐ 7. Other: _____

Applicant Must Provide:

☒ An initial or substitute computer readable form (CRF) copy of the "Sequence Listing".

☒ An initial or substitute paper copy of the "Sequence Listing", as well as an amendment directing its entry into the specification.

☒ A statement that the content of the paper and computer readable copies are the same and, where applicable, include no new matter, as required by 37 C.F.R. 1.821(e) or 1.821(f) or 1.821(g) or 1.825(b) or 1.825(d).

For questions regarding compliance to these requirements, please contact:

For Rules Interpretation, call (703) 308-4216
For CRF Submission Help, call (703) 308-4212
For PatentIn software help, call (703) 308-6856

PLEASE RETURN A COPY OF THIS NOTICE WITH YOUR RESPONSE